DOCUMENT RESUME

ED 266 685 FL 015 541

AUTHOR Evensen, Lars Sigfried
TITLE Testing Language Skills in the Norwegian Comprehensive School.
PUB DATE 85
NOTE 17p.; In: Language Testing in School. AFinLA Yearbook 1985. No. 41 (see FL 015 537). Document is in small print.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Ability; Communicative Competence (Languages); Curriculum Design; Data Interpretation; *Educational Change; Elementary Secondary Education; *English (Second Language); Foreign Countries; *Heterogeneous Grouping; Language Skills; *Language Tests; Public Schools; Second Language Instruction; Student Attitudes; Surveys; Teacher Attitudes; Testing; *Test Results
IDENTIFIERS Comprehensive Examinations; *Norway

ABSTRACT

Data from a national survey of students and teachers working within Norway's system of mixed-ability instruction in the elementary and early secondary schools suggest that, after ten years of implementation, the teaching and testing system is workable and that some aspects work relatively well. Instruction in English as a second language in a mixed-ability system has some problems that will be a continuing challenge for researchers and teacher trainers. When the curriculum was designed, the communicative aspect had not been worked out well, but it is anticipated that the revisions in the 1985 teaching guide will alter the previous non-communicative skills orientation to a more communicative one, especially in the direction of oral skills and cultural education. These changes are expected to lead to changes in testing. Some data interpretation and tables are included from both teacher and student responses and a national evaluation system. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as received from the person or organization orgnating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Tommola

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED266685

FL015541

TESTING LANGUAGE SKILLS IN THE FINNISH COMPREHENSIVE SCHOOL

Lars Sigfred Evensen (University of Trondheim)

1 Introduction

Finland is at a stage where streaming in PE is on its way out of the system and mixed-ability teaching and testing on its way in. Such a decision is going to have far-reaching effects on both teaching and testing. It is against this background that I would like to present selected data from the Norwegian system.

In Norway a curriculum based on the concept of mixed-ability teaching has been in effect since 1974. It grew out of a long period of pedagogical and even political struggle. In fact, the more fundamental questions behind the mixed-ability concept (eg. should we have one school for all?) have been central in Norwegian debate for more than one hundred years (Telhaug 1970:10). The reason for this is that one's choices on such issues will have effects that are not only linguistic and pedagogical, but also social. Education gives access to so many other things in society.

In the context of mixed-ability teaching testing is but one aspect. However, major pedagogical questions like individualisation have already been topics for separate conferences and seminars in Finland (cf. eg. seminar reports No.3 and No.11 from Nordiska språk- och informationscentret, Helsinki). Norwegian points of view have been presented by Frydenberg 1981a, 1981b and 1981c, by Stølen 1984a, 1984b and by Tangen 1984. Thus in my present talk it is natural to concentrate on questions of evaluation and assessment in English as a foreign language (EFL) in general and the exam system in particular.

As an applied linguist I feel that it is very important to supplement information about a system with info from the people who are actually working within the system. It is, after all, the wearer of the school, so to speak, who knows where it hurts most. For this reason I shall present selected results from a national survey carried out in 1981. The aim of this survey was to find out what a statistically representative national sample of students and teachers in grades 8-11 (student ages appr. 15-18) of Norwegian lower and upper secondary school perceive to be their problems in LL/LT in Norwegian as a first language and EFL (Evensen 1983).

BEST COPY AVAILABLE

In the survey evaluation stands as one out of a considerable number of problem areas. In the context of this conference I feel that it may be just as relevant to hear about such problems as to hear about any solutions that we might have to offer.

2 The Norwegian EFL curriculum and evaluation system

Teaching in Norwegian comprehensive school has since 1974 been based on the "Mønsterplan for grunnskolen" (Curriculum for primary and lower secondary education). The main educational innovation introduced by this plan was the step away from a fixed curriculum for all to a flexible teaching guide, where individual schools and teachers (at least in principle) are given considerable freedom to choose for themselves both what to teach and how to teach it. The testing system, however, is still largely centralized.

In the EFL teaching guide a major change consisted of moving away from explicit knowledge of language structure and metalanguage to an explicit emphasis on skills and a less explicit emphasis on language as a means of communication. The goal of EFL is stated in the following three parts (my translation):

> "The teaching of English is to aim at
> -giving the students practical language skill which may give them increased ability to establish contact, orally and in writing
> -forming as firm a basis as possible for continued learning of the subject
> -establishing such attitudes toward the language that the students develop an interest in increasing their skills and get an impulse to use the language."

Based on considerations of communicative needs, oral skills are given more emphasis than written skills. Furthermore, of the two written skills reading seems to be given priority before writing, which is partly given a secondary role (Simensen 1983):

> "The skill of written expression is also to be practiced, both as a support of training of the oral skills and as a means of expression in itself".

Given the fact that the teaching guide is based on general goals rather than specific objectives, with considerable local freedom within the system, one may suspect that teachers have problems in knowing what to evaluate, which criteria to use. In the national survey referred to, the teacher questionnaires contained the following question:

Do you think it is easy or difficult to know what to correct in this grade? (Put a ring around the figure below the appropriate response alternative).

very diffic.	quite diffic.	medium	quite easy	very easy	don't know
5	4	3	2	1	0

Figure 1. Translated item from EFL teacher's questionnaire.

The answers of 191 EFL teachers in grade 9 are displayed in Table 1 below. In the table answers are given as percentages for each response alternative. The frequencies are weighted to counterbalance biases introduced by disproportionate sampling design (cf. Evensen 1982a). Furthermore the frequencies are adjusted so that those repondents who answered 'don't know' (1.0 per cent of the total) or left the item unanswered (3.1 per cent of the total) are kept out of the computations.

Table 1. Selecting criteria for correcting student performance. EFL teachers' perceptions of difficulty in grade 9. Weighted and adjusted frequencies (n=191).

response alternatives					central tendency and dispersion	
very diffic. ---------- very easy						
5 %	4 %	3 %	2 %	1 %	Md	s (1)
02.9	13.5	49.5	27.3	06.8	2.821	.868

16.4 per cent, or about a sixth, of the teachers feel that knowing what to correct is very difficult or quite difficult. For these teachers the choice of evaluation criteria may be said to constitute a problem. For the great majority, however, the system seems to work quite well. If we use the response scale from five to one to compute an average (median) out of the answers given, the result is 2.821. This means that an "average answer" is a little toward the 'easy' end of the scale. The standard deviation adds the further information that there is some disagreement among teachers on the question; there is some degree of dispersion in the answers given.

The most reasonable conclusion one can draw from the data seems to be that, after seven years of experience with mixed-ability teaching, most teachers are in fact quite satisfied with the freedom (or burden) of choice built into the system. The table, in other words, seems

BEST COPY AVAILABLE

to show that local freedom is not as big a problem as one might have expected it to be for teachers with respect to knowing what to correct.

We may go on to ask what the students think of the teachers' evaluation. The student questionnaires contained an item of the type presented above where the students were asked to evaluate teacher demands, as demonstrated in teacher marking, on an ordinal scale going from 'much too high' to 'much too low'. The answers are presented in Table 2.

Table 2. Marking. EFL students' perceptions of teacher demands in grade 9. Weighted and adjusted frequencies (n=410).

response alternatives: demands are---					central tendency and dispersion	
much too great				much too small	Md	s
5 %	4 %	3 %	2 %	1 %		
12.7	30.9	49.7	06.8	00.0	3.371	.801

It seems that the students are not totally satisfied with the state of affairs. 43.6 per cent feel that teacher demands are 'much too great' or 'a little too great'. On the other hand nobody seems to feel that the teachers are much too lenient. 6.8 per cent think that the teachers are a little too lenient. The median reflects this distribution. A result of 3.371 is toward the 'too great' part of the scale. Still, the median is so close to the neutral mid category that this problem may not in any sense be said to be an important one.

3 The Norwegian EFL exam

In this section I shall present some information about the Norwegian EFL exam at the end of comprehensive school. In the presentation I shall focus on the correspondence between the goals and the exam system. I shall also consider potential backwash effects on everyday teaching. It is a well-known fact that the exam system influences practical teaching. Simensen (1983:217) has formulated this principle very strongly; if there is a discrepancy between goals and evaluation, evaluation usually wins.

BEST COPY AVAILABL

I have chosen to use teacher perceptions of correspondence as my starting point. In the questionnaire teachers were asked to give their views as to the degree of correspondence between the goals and the EFL exam. I find the results very interesting.

Table 3. Teacher perceptions of the correspondence between the official goals and the EFL exam at the end of comprehensive school. Weighted and adjusted frequencies (n=191).

response alternatives					central tendency and dispersion	
very good corr.	----------	----------	----------	very bad corr.	Md	s
5 %	4 %	3 %	2 %	1 %		
00.0	03.6	49.8	36.8	09.9	2.568	.722

It seems clear that we deal here with a real problem of some intensity. None of the teachers think that the correspondence is 'very good', and only 3.6 per cent feel that it is 'quite good'. On the other hand almost ten per cent think that the correspondence is 'very bad'. The median is as low as 2.568 and the low standard deviation (.722) indicates a higher degree of concensus among the teachers on this question than on previous ones. In an effort to understand this problem I shall first present some background information about the exam system and then supplement this material with more detailed survey data.

English is the only foreign language with a final written exam in Norway. The test takes place at the end of ninth grade and is allotted four hours (+ fifteen minutes for practicalities). It officially consists of at least four out of a list of eight subtests presented in the Handbook for the school, part III (Håndbok for skolen, del III):

-comprehension questions on a given text
-fill-in-the-blank (including cloze)
-completion of a text where the beginning or end has been removed
-free comments on a text
-essay based on key words
-picture-based essay
-essay on a given topic

According to the regulations each student sits for at least one of the following subjects-

BEST COPY AVAILABLE

-Norwegian as a first language
-Mathematics
-EFL

In practice each student sits for only one of these, which means that all students in approximately one third of the country's nineteen counties sit for the EFL exam every year (cf. Table 5 below). The particular counties are selected anew each year according to a confidential system.

Each student's performance is evaluated by a team of two external examiners each evaluating about 150 student performances. Examiners work independently at first, then change student answers and at the end discuss borderline cases.

Since 1982 the guiding principle behind evaluation criteria and marking has been one called "adapted goal-referenced testing" (cf. Norsk Skole Nos. 10/11 and 12/13 1982 and 4 1984). This principle may be seen as an attempt to combine goal-referenced, criterion-referenced and norm-referenced testing.

The main basis for marking is the examiners' overall impression of each student's performance. In doing this the examiners are told to relate their judgement not just to linguistic aspects, but also to the overall goals of the school system, as these are stated in the teaching guide. Apart from a general description of intentions the examiners are not given many details as to how to go about implementing the principle in actual marking.

More detailed evaluation criteria are distributed every year to the examiners and later published in the series Evaluering i grunnskolen (Evaluation in the comprehensive school) issued by The National Council for Primary and Lower-secondary Education.

For the more closed subtests in the written exam, some of which are discrete-point, a number of points are given for each subtest (implying a relative ranking of them as to importance) and a total sum for all subtests is suggested for each mark. In certain cases, a graded system is used for the allotment of points to each individual answer, where the top score is reserved for those answers that are both formally and functionally correct (eg. a question with do-periphrasis). A medium score is given for understandable but formally incorrect answers (eg. "How long do the journey takes?"), and no points are given for answers that are both formally and functionally incorrect. Thus communicative principles have to some extent entered into the criterion-referenced part of the marking.

For the essays several broad criteria are suggested. The essays in the upper two bands of the five-point marking scale (see below) should be "good" as to coherence (2), diction and vocabulary, having few errors of orthography and verb inflection. Furthermore the essays have to be in accordance with the wording of the task given.

The essays in the mid group are accepted with somewhat lower standard as to coherence, diction, vocabulary, errors of orthography and verb inflection; some naive formulations and a relatively short answer. The criteria for the low group are more indirect; a very short essay will not receive good marks and an otherwise medium essay should not be lowered more than one step in the marking if the student has not written the essay part of the exam. For the lowest mark the criterion is communicative; the essay would not be understood if performed orally.

By way of contrast, the oral exam consists of a combination taken out of a list of four subtests (Håndbok for skolen, del III):

-reading of a passage from a known text followed by conversation about the text from which the passage is taken
-free conversation about everyday topics
-reading of a short, unknown text followed by comprehension questions
-conversation about an unknown text read aloud by the teacher (the internal examiner)

Certain students from two counties are selected to sit for the oral exam in English every year. This means that the proportion of students sitting for the oral is much lower than the proportion sitting for the written exam (cf. Table 5 below).

There is one external examiner for each student, and each team of examiners evaluates the performance of some 20 students. In the marking regulations the emphasis is explicitly on the students' ability to carry on a conversation, which is more important than her/his pronunciation and intonation. The regulations do, here as well, have a communicative element.

Statistics as to the distribution of subtests are only available for written exams. In the following section I shall concentrate on the written exam, supplementing some statistics and survey data that are relevant also for the oral.

The norm-referenced part of the evaluation criteria is quite indirect. National standardized tests in the structuralist-psychometric tradition (Spolsky) have since 1962-63 been offered on an optional basis as one

element guiding the teachers' achievement marking. This system is used by 80-90 per cent of the schools (OMI report No. 20, p. 18). Although no norm is given for the national exam, it may be assumed that these tests have long-range effects on individual teacher marking both because of their long tradition and because of their widespread use (for further evidence, see Table 5 below). The weight of the test is said to equal one double-lesson school test. The marks are to be normally distributed (5 is the top mark - "particularly good"- and Lq is the bottom mark - "little of positive value"):

5	M	G	Mq	Lq
++	+	/	-	--
4%	24%	48%	24%	4%

Let us look at the characteristics of the written EFL exam during the last five years. The distribution of subtests is presented in Table 4.

Table 4. Distribution of subtests in recent EFL written exams.

	1980	'81	'82	'83	'84
-open-ended comprehension questions	X	X	X	X	-
-true/false items	X	-	-	-	X
-text-based questions/ answers	X	X	X	X	XX
-story/dialogue completions	-	-	-	-	X
-vocabulary test	X	-	X	XX	X
-cloze/ fill-in-the-blank	X	-	X	X	XX
-transformations and paraphrases	X	XX	-	-	-
-sentence completions	-	-	X	-	-
-text-based free comments	-	-	-	-	-
-expression of selected language functions	-	X	-	-	-
-letters/ postcards	X	-	X	X	-
-essay (1 out of 4-6 tasks)	X	X	X	X	X

The distribution of subtests clearly indicates that the average number of subtests is larger than four. Furthermore, several subtests that are not included in the official list appear in the actual exams (eg. transformations and sentence completion) whereas one subtest is not used at all in the period studied (free comments on a text). Two test types have dominated the exams in the period in the sense that they have been used every year - making questions or answers to a text and essay writing.

The frequency of the former subtest calls for a comment. The most likely explanation for this perhaps unexpected pattern is suggested in a critical study by Simensen (1979). On the basis of a longitudinal study of EFL curriculum development and exams she concluded that several aspects of the exam tradition might be questioned when related to the official curricula. One such aspect is the preponderance of items where the students are only trained to answer given questions when in real-life communication they also need to be able to ask questions. An inspection of the exam tasks in the period under study shows that it is in fact this aspect that is emphasized; this subtest in all cases following Simensen's report asks the student to make questions that are relevant to a given text or situation.

In Simensen's (1979) study the suspicion that some subtests may not discriminate well was also expressed. She furthermore suggested that the use of meta-language and discrete-point tests of grammar should be abolished. In the period under study it appears that there has been an increase in the number of subtests where two versions differing in difficulty (eg. in that one subtest is guided where a similar subtest is not or that the marking of one subtest takes accuracy into account where the other does not) are used.

In the period after Simensen's study no subtests have relied on metalinguistic knowledge. However to some extent discrete-point tests of grammar have still been used, but not since 1982. It may be justified to conclude that one positive aspect of the Norwegian system is its ability to take outside criticism into consideration.

One point that is not taken into consideration, however, is Simensen's criticism of the written language bias in the system. This bias demonstrates itself in the distribution of written and oral exams in the period under study. Table 5 presents both this distribution and the distribution of marks in the four-year period between 1980 and 1983 (the data from 1984 are not yet available).

Table 5. Marks, norms and exam results in ENG, 1980-83.

Marks	S ++	M +	G /	Mg -	Lg --
Norms	4%	24%	48%	24%	4%
Achievement marks:					
1980 (n=62.648)	2.91	31.11	40.46	23.40	2.12
1981 (n=64.329)	3.26	30.87	40.72	23.02	2.12
1982 (n=65.895)	3.24	31.32	40.51	22.75	2.18
1983 (n=65.219)	3.18	31.72	40.70	22.35	2.05

Written exams: (all students in seven counties)					
1980 (n=25.011 =38.2%)	3.40	22.73	43.51	25.75	4.61
1981 (n=26.975 =40.2%)	3.99	23.73	44.18	23.80	4.30
1982 (n=22.331 =35.6%)	2.95	21.57	44.67	25.70	5.10
1983 (n=21.657 =31.9%)	2.96	23.20	44.96	24.76	4.11
Oral exams: (some students in two counties)					
1980 (n=1.263 =1.9%)	8.95	41.41	34.60	14.33	0.71
1981 (n=2.043 =3.0%)	6.22	48.21	33.72	11.50	0.34
1982 (n=1.330 =1.9%)	6.77	40.98	36.99	14.96	0.30
1983 (n=1.264 =1.9%	6.65	40.03	40.11	12.82	0.40

The table makes it abundantly clear that the written exam dominates when seen in relation to the oral exam. Whereas between 30 and 40 per cent of the students sit for the written exam every year, only two to three per cent of the students sit for the oral. The individual student may thus for all practical purposes count on not having to sit for the oral. This backwash effect is serious when related to the official curriculum's emphasis on oral communication.

The distribution of marks for both achievement and written exam is surprisingly close to the normal distribution suggested in the regulations, particularly so for the written exam. In the achievement marks there is a weak tendency to avoid using the extreme marks. There is also a weak tendency to skew the marks toward the positive end of the scale, and the curve is generally somewhat flatter than officially suggested. Still there seems to be good reason to assume a relatively strong indirect effect of the standardized tests even if a normal distribution is no longer presupposed in the adapted goal-referenced evaluation. For the oral exam, however, the effect is not strong. Here about half of the students are found in the upper two bands.

On the basis of this material I would like to return to the question of teacher discontent with the goals-exam correspondence. Additional data are presented in Table 6. Here the teachers were asked if they think

BEST COPY AVAILABLE

that each of a number of areas is given the right emphasis in the present exam system.

Table 6. Relative emphasis on skills and knowledge in the present exam system. EFL teachers' views. Weighted and adjusted frequencies (n=191).

	alternatives					central tendency and dispersion	
	too much emphasis 5 %	4 %	3 %	2 %	too little emphasis 1 %	Md	s
-written skills	19.6	34.4	34.2	11.2	00.6	3.617	.969
-oral skills	00.0	01.0	31.9	30.9	36.3	1.944	.890
-cultural knowledge	00.0	05.5	39.4	43.2	11.8	2.382	.767
-linguistic knowledge	00.7	04.2	70.4	20.6	04.1	2.868	.624

There is a major discrepancy between the evaluation of written and oral skills in the teacher material. Written skills are generally seen as having too much emphasis in the EFL exam, whereas oral skills are seen as having too little emphasis. In fact more than a third of the teachers feel that oral skills are given far too little emphasis. The median is as low as 1.996, indicating that this problem has considerable intensity (3). At this point it seems reasonable to suspect that at least one major cause of teacher discontent has been located.

The table also indicates that the cultural aspects of EFL are given too little weight. As to linguistic knowledge the neutral mid category is chosen by as many as 70.4 per cent of the teachers. In the Norwegian curriculum linguistic knowledge is deemphasized, and in recent exams it has not been tested in an explicit manner. It is interesting to note that so many teachers are satisfied with the exam in this respect. The standard deviation for this variable is lower than for any other variable we have analysed so far. In other words there is both considerable satisfaction and agreement among the teachers on this question.

We have already to some extent considered the backwash effect of the exam. In a broader perspective different subtests may be evaluated as to their motivational effect on students. In the national survey the student respondents were asked about likes and dislikes as to selected test types.

BEST COPY AVAILABLE

Table 7. Written EFL test forms ranked according to grade 9 students' preferences. Weighted and adjusted frequences (n=410).

	alternatives					central tendency and dispersion	
	like				dislike		
	5 %	4 %	3 %	2 %	1 %	Md	s
comprehension questions	12.4	41.3	30.6	12.5	03.2	3.590	.971
fill-in-the-blank grammar questions	15.3	33.8	40.9	08.4	01.6	3.477	.906
	08.9	12.1	32.4	31.5	15.1	2.604	1.139
free essay	24.3	23.4	24.6	16.3	11.4	3.407	1.312
picure-based essay	08.9	23.7	36.6	18.7	12.1	3.025	1.126
narrative essay	09.3	20.2	27.6	24.6	18.7	2.756	1.231
expository essay	04.0	11.3	30.8	32.4	21.5	2.380	1.070
descriptive essay	04.4	17.9	21.7	35.6	20.4	2.331	1.111

From the students' point of view the essay part is the least popular. Discrete-point grammar questions are the only exception to this pattern. Both comprehension questions to a given text and subtests of the fill-in-the-blank type (including cloze) are generally well thought of.

Essays are not popular. However, the large dispersion both between and within essay types indicate considerable disagreement among the students. Free essays (where the individual student chooses her own topic) are generally well thought of. There are also groups of students who like other essay types.

In order to conclude the data section I would like to put the question of mixed-ability testing within the larger context of mixed-ability teaching. In a conference of testing we might easily forget that there are other and perhaps even more problematic areas to consider. The teacher questionnaires contained a section where the teachers were asked to evaluate their own teacher training with respect to a number of task areas relevant to their jobs. The results on a selected number of variables may shed some light on the question of evaluation versus other areas of teaching in a mixed-ability context.

Table 8. EFL teachers' evaluation of their own teacher training. Weighted and adjusted frequencies (n=1911).

	alternatives					central tendency and dispersion	
own training as to:	very good 5 %	4 %	3 %	2 %	very bad 1 %	M	s
knowledge	38.8	38.6	16.8	05.8	00.0	4.210	.805
skills	29.4	41.2	24.9	04.4	00.0	4.001	.851
evaluation	05.9	09.9	36.0	22.9	25.2	2.983	1.147
differentiation	00.0	05.9	18.6	38.2	37.3	1.832	.892

It is evident that most teachers in the present sample are very well satisfied with both the skills and knowledge aspects of their own training. With the more pedagogical aspects of their training, however, the picture is different. Furthermore there is a clear difference between evaluation and differentiation in the data. Evaluation is a relatively mild problem after all in this context. Differentiation is a much more serious problem. This result is just one of several in the total survey material which suggests that from the teachers' point-of-view mixed-ability teaching is a far more serious problem in EFL than mixed-ability evaluation.

Concluding remarks

In the present paper I have tried to combine data from a national evaluation system with data from people working in the system. I think it is a fair conclusion after ten years of mixed-ability teaching and testing that such systems are quite possible to implement. Some aspects of these systems, as noted in this paper, actually seem to work well.

It also seems justified to conclude that when we choose to settle for a mixed-ability system we are not taking an easy way out as far as the organization of EFL is concerned. The system has its problems, particulary in practical teaching. These are going to be a continuing challenge for researchers and teacher trainers for a long time after the implementation of mixed-ability teaching and testing.

Haastrup (this seminar) has questionned the construct validity of the Danish oral exam in relation to the notion of communicative competence. This reservation is probably justified both for the oral and written exams in Norway. Here, the central problem seems to be

establishing both theoretically well-founded and precise definitions of central but vague terms like eg. 'coherence' (cf. footnote 2).

In Norway the communicative paradigm was not well established at the time when its curriculum was designed. It is to be expected that the 1985 revision of the teaching guide will change the basic non-communicative skills orientation into a more communicative one. This is hopefully going to lead to changes also in testing. In this connection, the material I have presented seems to call first for a reorientation in the direction of oral skills. It may also be hoped that the cultural aspects of foreign language teaching and learning will be emphasized more. Communicative competence is, after all, not simply a question of skill.

Notes

1 The statistics were produced by an old version of SPSS which does not compute percentiles.

2 It is significant to note that this term is not treated systematically in teacher training, teaching materials or actual teaching. At the University of Trondheim's Department of Applied Linguistics research is in progress to investigate students' written performance at the discourse-level. The research is based on the Trondheim Corpus of Applied Linguistics (Evensen 1982b), which consists of compositions written by the 2295 students who took part in the national survey. A central aim of the research is to carry out performance analyses to find correspondences between (non-) use of discourse-level features and holistic teacher evaluation in different grades. Preliminary results from exploratory studies are reported in (Evensen (in press) and in Evensen (forthc.).

3 Considering both the well-known error of central tendency (cf. eg. Oppenheim 1966) and the fact that the measuring scale has only five steps, a median of 1.994 is very low indeed.

Bibliography

Differensieringsfrågor i språkundervisningen. Seminarierapport nr. 11. 1983. Helsinki: Nordiska språk- och informationscentret.

Evaluering i grunnskolen. 1980-1983. Grunnskolerådet. Oslo: Grøndahl.

Enkvist, N.E., ed. in press. Composition and coherence: a symposium. Publications of the Research Institute of the Abo Akademi Foundation. Abo: Abo Akademi.

Evensen, L.S. 1982a. "Forskning for morgendagens fremmedspråksundervisning." In the report Lærere og skoleforskning. Oslo: NAVF.

- 1982b. "Data om språkundervisning". In Humanistiska Data, Vol. 9, No. 2, 4-11.

- 1984. "Difficulty in EFL: a national survey." In Ringbom & Rissanen (1984:349-358).

- in press. "Discourse-level interlanguage studies". To appear in Enkvist (in press).

- forthc. "Connecting L1 and FL in discourse-level performance analysis." Paper presented at the 20th International conference of contrastive linguistics, Posnan, Poland, Dec. 13-18, 1984.

Frydenberg, R. 1981a. "Differensiering i språkundervisningen." In Undervisningas differentiering.... pp. 26-31.

- 1981b. "Målen och hur de uppnåtts samt utvecklingstrender i Norge. Undervisningens.... pp. 56-58.

Frydenberg, R. 1981c. "Rollspel i undervisningen i främmande språk." In Undervisningens... pp. 152-194.

Håndbok for skolen, del III. Kirke- og undervisningsdepartementet. Oslo.

Kunnskapsformidling i en skole for alle: innhold og nivå. OMI-rapport nr. 22. Oslo: Grunnskolerådet.

Norsk Skole, 1982,1984. Oslo: Kirke- og undervisningsdepartementet.

Oller, J. 1979. Language tests at school. London: Longman.

Oppenheim, A.N. 1966. Questionnaire design and attitude measurement. London: Heinemann.

Ringbom, H. & Rissanen, M., eds. 1984. Proceedings from the Second Nordic conference for English studies. Publications of the Research Institute of the Abo Akademi Foundation, No. 92. Abo: Abo Akademi.

Simensen, A. M. 1979. Evaluering i engelsk skriftlig ved avgangseksamen i 9-årig skole: en beskrivelse og en analyse. Report. Trondheim: University of Trondheim.

- 1983. "Intensjon og virkelighet i engelsk som fremmedspråk." In Kunnskaps-formidling.... pp.212-232.

Stølen,L. 1984a. "Idéer og forslag til lærernes grunn utdanning og videreutdanning." In Differensieringsfrågor....pp. 74-77.

- 1984b. "Resultat av forsøksvirksomhet innen undervisningens differensie ring i Norge." In Differentieringsfrågor.... pp.78-79.

Tangen, K. 1984. "Ressurser til grunnskolen i Norge." In Differentieringsfrågor....pp. 80-107.

Telhaug, A.O. 1970. Den ni-årige skolen og differensieringsproblemet. Oslo: Lærerstudentenes Forlag.

 BEST COPY AVAILABLE

Undervisningsdifferentiering i främmande språk på grundskolans högstadium. Seminarierapport nr. 11. 1984. Helsinki: Nordiska språk- och informationscentret.

Utviklingstendenser i elevprestasjoner på ungdomstrinnet. OMI-rapport nr. 20. 1983. Oslo: Universitetsforlaget.

BEST COPY AVAILABLE